A taxonomy of prompt modifiers for text-to-image generation

Year: 2022 | Citations: 139 | Authors: J. Oppenlaender

Abstract

ABSTRACT Text-guided synthesis of images has become enormously popular and online communities dedicated to text-to-image generation and art generated with Artificial Intelligence (AI) have emerged. While deep generative models can synthesise high-quality images and artworks from simple descriptive text prompts, practitioners of text-to-image generation typically seek to control the generative model's output by adding short key phrases ('modifiers') to the prompt. This paper identifies six types of prompt modifiers used by practitioners in the online text-to-image community based on a 3-month ethnographic study. The novel taxonomy of prompt modifiers provides researchers a conceptual starting point for investigating the practice of text-to-image generation, but may also help practitioners of AI generated art improve their images. We further outline how prompt modifiers are applied in the practice of 'prompt engineering.' and discuss research opportunities of this novel creative practice in the field of Human–Computer Interaction (HCI). The paper concludes with a discussion of broader implications of prompt engineering from the perspective of Human-AI Interaction (HAI) in future applications beyond the use case of text-to-image generation and AI generated art.